

1652

RAW SEQUENCE LISTING                      DATE: 10/19/2000  
 PATENT APPLICATION: US/09/308,207              TIME: 08:34:10

Input Set : A:\ES.txt  
 Output Set: N:\CRF3\10192000\I308207.raw

## SEQUENCE LISTING

5 (1) GENERAL INFORMATION:  
 7     (i) APPLICANT: MARIA DIAZ-TORRES ET AL.  
 9     (ii) TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
 10                                      PRODUCTION OF 1,3 PROPANEDIOL  
 12     (iii) NUMBER OF SEQUENCES: 68  
 14     (iv) CORRESPONDENCE ADDRESS:  
 15             (A) ADDRESSEE: Genencor International, Inc.  
 16             (B) STREET: 4 Cambridge Place  
 17                                      1870 South Winton road  
 18             (C) CITY: Rochester  
 19             (D) STATE: NY  
 20             (E) COUNTRY: U.S.A  
 21             (F) ZIP: 14618  
 23     (v) COMPUTER READABLE FORM:  
 24             (A) MEDIUM TYPE: Diskette  
 25             (B) COMPUTER: IBM Compatible  
 26             (C) OPERATING SYSTEM: Windows  
 27             (D) SOFTWARE: FastSEQ for Windows Version 2.0b  
 29     (vi) CURRENT APPLICATION DATA:  
 C--> 30             (A) APPLICATION NUMBER: US/09/308,207.  
 C--> 31             (B) FILING DATE: 13-May-1999  
 32             (C) CLASSIFICATION:  
 34     (vii) PRIOR APPLICATION DATA:  
 C--> 35             (A) APPLICATION NUMBER: 60/030,601  
 C--> 36             (B) FILING DATE: 13-NOV-1996  
 40     (viii) ATTORNEY/AGENT INFORMATION:  
 41             (A) NAME: Glaister, Debra  
 42             (B) REGISTRATION NUMBER: 33,888  
 43             (C) REFERENCE/DOCKET NUMBER: GC 369-2  
 45     (ix) TELECOMMUNICATION INFORMATION:  
 46             (A) TELEPHONE: 650-864-7620  
 47             (B) TELEFAX: 650-845-6504  
 48             (C) TELEX:  
 51 (2) INFORMATION FOR SEQ ID NO: 1:  
 53     (i) SEQUENCE CHARACTERISTICS:  
 54             (A) LENGTH: 1668 base pairs  
 55             (B) TYPE: nucleic acid  
 56             (C) STRANDEDNESS: single  
 57             (D) TOPOLOGY: linear  
 59     (ii) MOLECULE TYPE: DNA (genomic)  
 61     (iii) HYPOTHETICAL: NO  
 63     (iv) ANTI-SENSE: NO  
 65     (vi) ORIGINAL SOURCE:  
 66             (A) ORGANISM: DHAB1  
 C--> 68     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 70 ATGAAAAGAT CAAAACGATT TGCAGTACTG GCCCAGCGCC CCGTCAATCA GGACGGGCTG   60

ENTERED

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/308,207  
 DATE: 10/19/2000  
 TIME: 08:34:10

Input Set : A:\ES.txt  
 Output Set : N:\CRF3\10192000\I308207.raw

```

72 ATTGGCGAGT GGCCTGAAGA GGGGCTGATC GCCATGGACA GCCCCTTTGA CCCGGTCTCT 120
74 TCAGTAAAAG TGGACAACGG TCTGATCGTC GAACTGGACG GCAAACGCCG GGACCAGTTT 180
76 GACATGATCG ACCGATTATG CGCCGATTAC GCGATCAACG TTGAGCGCAC AGAGCAGGCA 240
78 ATGCGCCTGG AGGCGGTGGA AATAGCCCGT ATGCTGGTGG ATATTCACGT CAGCCGGGAG 300
80 GAGATCATTG CCATCACTAC CGCCATCAGC CCGGCCAAAG CGGTCGAGGT GATGGCGCAG 360
82 ATGAACGTGG TGGAGATGAT GATGGCGCTG CAGAAGATGC GTGCCCCCGG GACCCCTCC 420
84 AACCAGTGCC ACGTCACCAA TCTCAAAGAT AATCCGGTGC AGATTGCCGC TGACGCCGCC 480
86 GAGGCCGGGA TCCGCGGCTT CTCAGAACAG GAGACCACGG TCGGTATCGC GCGCTACGCG 540
88 CCGTTTAACG CCCTGGCGCT GTTGGTCGGT TCGCAGTGC GCGCCCCGG CGTGTGACG 600
90 CAGTGCTCGG TGGAGAGGCG CACCGAGCTG GAGCTGGGCA TCGTGGCTT AACCAGCTAC 660
92 GCCGAGACGG TGTCGGTCTA CGGCACCGAA GCGGTATTTA CCGACGGCGA TGATACGCCG 720
94 TGGTCAAAGG CGTTCCTCGC CTCGGCCTAC GCCTCCCGCG GGTGAAAAAT GCGCTACACC 780
96 TCCGGCACCG GATCCGAAGC GCTGATGGGC TATTCGAGA GCAAGTCGAT GCTCTACCTC 840
98 GAATCGCGCT GCATCTTCAT TACTAAAGC GCGGGGTTT AGGACTGCA AACGGCGCG 900
100 GTGAGCTGTA TCGGCATGAC CGGCGCTGTG CCGTCGGGCA TTCGGGCGGT GCTGGCGGAA 960
102 AACCTGATCG CCTCTATGCT CGAECTCGAA GTGGCGTCCG CCAACGACCA GACTTTCTCC 1020
104 CACTCGGATA TTCGCGCAC CGCGCGCACC CTGATGCAGA TGCTGCCGGG CACCGACTTT 1080
106 ATTTTCTCCG GCTACAGCGC GGTGCCGAAC TACGACAACA TGTTGCCGG CTGCACTTC 1140
108 GATGCGGAAG ATTTTGATGA TTACAACATC CTGCAGCGTG ACCTGATGGT TGACGGCGGC 1200
110 CTGCGTCCGG TGACCGAGCG GGAACCAATT GCCATTGCGC AGAAAGCGGC GCGGGCGATC 1260
112 CAGGCGGTTT TCCGCGAGCT GGGGCTGCCG CCAATCGCCG ACGAGGAGGT GGAGGCCGCC 1320
114 ACCTACGCGC ACGGCAGCAA CGAGATGCGG CCGCGTAACG TGGTGGAGGA TCTGAGTGCG 1380
116 GTGGAAGAGA TGATGAAGCG CAACATCACC GGCTCGATA TTGTCGGCGC GCTGAGCCGC 1440
118 AGCGGCTTTG AGGATATCGC CAGCAATATT CTCAATATG TCGCCAGCGG GGTCAACGGC 1500
120 GATTACCTGC AGACCTCGGC CATTCCTGAT CGGCAGTTTC AGGTGGTGAG TGCGGTCAAC 1560
122 GACATCAATG ACTATCAGGG GCCGGGCACC GGCTATCGCA TCTCTGCCGA ACGCTGGGCG 1620
124 GAGATCAAAA ATATTCCGGG CGTGGTTTCA CCCGACACCA TTGAATAA 1668
126 (2) INFORMATION FOR SEQ ID NO: 2:
128 (i) SEQUENCE CHARACTERISTICS:
129 (A) LENGTH: 585 base pairs
130 (B) TYPE: nucleic acid
131 (C) STRANDEDNESS: single
132 (D) TOPOLOGY: linear
134 (ii) MOLECULE TYPE: DNA (genomic)
136 (vi) ORIGINAL SOURCE:
137 (A) ORGANISM: DHAB2
C--> 139 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
141 GTGCAACAGA CAACCCAAAT TCAGCCCTCT TTTACCTGA AAACCCGCGA GGGCGGGTA 60
143 GCTTCTGCCG ATGAACGCGC CGATGAAGTG GTGATCGGCG TCGGCCCTGC CTCGATAAAA 120
145 CACCAGCATC ACACTCTGAT CGATATGCCC CATGGCGCGA TCCTCAAAGA GCTGATTGCC 180
147 GGGGTGGAAG AAGAGGGGCT TCACGCCCGG GTGGTGCGCA TTCTGCGCAC GTCCGACGTC 240
149 TCCTTTATGG CCTGGGATGC GGCCAACCTG AGCGGCTCGG GATCGGCAT CGGTATCCAG 300
151 TCGAAGGGGA CCACGGTCAT CCATCAGCGC GATCTGCTGC CGCTCAGCAA CCTGGAGCTG 360
153 TTCTCCCAAG CCGCGTGCT GACGCTGGAG ACCTACCGGC AGATTGGCAA AAACGCTGCG 420
155 CGCTATGCGC GCAAAGAGTC ACCTTCGCCG GTGCCGGTGG TGAACGATCA GATGGTGCGG 480
157 CCGAAATTTA TGGCCAAAGC CGCGCTATTT CATATCAAAG AGACCAAACA TGTGGTGACG 540
159 GACGCCGAGC CCGTCAACCT GCACATCGAC TTAGTAAGGG AGTGA 585
161 (2) INFORMATION FOR SEQ ID NO: 3:
163 (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING                      DATE: 10/19/2000  
 PATENT APPLICATION: US/09/308,207        TIME: 08:34:10

Input Set : A:\ES.txt  
 Output Set: N:\CRF3\10192000\I308207.raw

```

164      (A) LENGTH: 426 base pairs
165      (B) TYPE: nucleic acid
166      (C) STRANDEDNESS: single
167      (D) TOPOLOGY: linear
169      (ii) MOLECULE TYPE: DNA (genomic)
171      (vi) ORIGINAL SOURCE:
172      (A) ORGANISM: DHAB3
C--> 174      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
176 ATGAGCGAGA AAACCATGCG CGTGCAGGAT TATCCGTTAG CCACCCGCTG CCCGGAGCAT    60
178 ATCCTGACGC CTACCGGCAA ACCATTGACC GATATTACCC TCGAGAAGGT GCTCTCTGGC    120
180 GAGGTGGGCC CGCAGGATGT GCGGATCTCC CGCCAGACCC TTGAGTACCA GCGCGAGATT    180
182 GCCGAGCAGA TGCAGCGCCA TGCAGTGGCG CGCAATTTCG GCCGCGCGGC GGAGCTTATC    240
184 GCCATTCTCG ACGAGCGCAT TCTGGCTATC TATAACGCGC TCGCGCCGTT CCGCTCCTCG    300
186 CAGGCGGAGC TGCTGGCGAT CGCCGACGAG CTGGAGCACA CCTGGCATGC GACAGTGAAT    360
188 GCCGCTTTTG TCCGGGAGTC GCGGAAGTG TATCAGCAGC GGCATAAGCT GCGTAAAGGA    420
190 AGCTAA                                           426
192 (2) INFORMATION FOR SEQ ID NO: 4:
194      (i) SEQUENCE CHARACTERISTICS:
195      (A) LENGTH: 1164 base pairs
196      (B) TYPE: nucleic acid
197      (C) STRANDEDNESS: single
198      (D) TOPOLOGY: linear
200      (ii) MOLECULE TYPE: DNA (genomic)
202      (vi) ORIGINAL SOURCE:
203      (A) ORGANISM: DHAT
C--> 205      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
207 ATGAGCTATC GTATGTTGA TTATCTGGTG CCAAACGTTA ACTTTTTTGG CCCCAACGCC    60
209 ATTTCCGTAG TCGGCGAAGC CTGCCAGCTG CTGGGGGGGA AAAAAGCCCT GCTGTCACC    120
211 GACAAAGGCC TCGCGGCAAT TAAAGATGGC GCGGTGGACA AAACCCCTGCA TTATCTGCGG    180
213 GAGGCCGGGA TCGAGGTGGC GATCTTTGAC GCGGTCGAGC CGAACCCGAA AGACACCAAC    240
215 GTGCGCGACG GCCTCGCCGT GTTTCGCCGC GAACAGTGCG ACATCATCGT CACCGTGGGC    300
217 GCGGCGAGCC CGCAGGATTG CGGCAAGGC ATCGGCATCG CCGCCACCCA TGAGGGCGAT    360
219 CTGTACCAAGT ATGCCGGAAT CGAGACCTTG ACCAACCOCG TGCCGCCTAT CGTCGCGGTC    420
221 AATACCACCG CCGGCACCGC CAGCGAGGTC ACCCGCCACT GCGTCCTGAC CAACACCGAA    480
223 ACCAAAGTGA AGTTTGTGAT CGTCAGCTGG CGCAAAGTGC CGTCGGTCTC TATCAACGAT    540
225 CCACTGCTGA TGATCGGTAA ACGGCGCGCC CTGACCGCGG CGACCGGGAT GGATGCCCTG    600
227 ACCCACGCGC TAGAGGCCTA TATCTCCAAA GACGCTAACC CCGTGACGGA CGCCGCCGCC    660
229 ATGCAGGCGA TCCGCCTCAT CGCCCGCAAC CTGCGCCAGG CCGTGGCCCT CGGCAGCAAT    720
231 CTGCAGGCGC GGGAAAACAT GGCCATATGCT TCTCTGCTGG CCGGGATGGC TTTCAATAAC    780
233 GCCAACCTCG GCTACGTGCA CGCCATGGCG CACCAGCTGG GCGGCCTGTA CGACATGCCG    840
235 CACGGCGTGG CCAACGCTGT CCTGCTGCCG CATGTGGCGC GCTACAACCT GATCGCCAAC    900
237 CCGGAGAAAT TCGCCGATAT CGCTGAACTG ATGGGCGAAA ATATACCGG ACTGTCCACT    960
239 CTCGACGCGG CGGAAAAAGC CATCGCCGCT ATCACGCGTC TGTCGATGGA TATCGGTATT    1020
241 CCGCAGCATC TCGCGCATCT GGGGGTAAAA GAGGCCGACT TCCCTACAT GGCGGAGATG    1080
243 GCTCTAAAAG ACGCAATGC GTTCTCGAAC CCGCGTAAAG GCAACGAGCA GGAGATTGCC    1140
245 GCGATTTTCC GCCAGGCATT CTGA                                           1164
247 (2) INFORMATION FOR SEQ ID NO: 5:
249      (i) SEQUENCE CHARACTERISTICS:
250      (A) LENGTH: 1380 base pairs

```

## RAW SEQUENCE LISTING

DATE: 10/19/2000

PATENT APPLICATION: US/09/308,207

TIME: 08:34:10

Input Set : A:\ES.txt

Output Set : N:\CRF3\10192000\I308207.raw

```

251      (B) TYPE: nucleic acid
252      (C) STRANDEDNESS: single
253      (D) TOPOLOGY: linear
255      (ii) MOLECULE TYPE: DNA (genomic)
257      (vi) ORIGINAL SOURCE:
258      (A) ORGANISM: GPD1
C--> 260      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
262 CTTTAATTTT CTTTATCTT ACTCTCCTAC ATAAGACATC AAGAAACAAT TGTATATGTG      60
264 ACACCCCCC CCTCCACAAA CACAAATATT GATAATATAA AGATGTCTGC TGCTGCTGAT      120
266 AGATTAAACT TAACTTCCCG CCACTTGAAT GCTGGTAGAA AGAGAAGTTC CTCTTCTGTT      180
268 TCTTTGAAGG CTGCCGAAAA GCCTTTCAAG GTTACTGTGA TTGGATCTGG TAACTGGGGT      240
270 ACTACTATTG CCAAGGTGGT TGCCGAAAAA TGTAAGGGAT ACCCAGAAGT TTTGCGTCCA      300
272 ATAGTACAAA TGTGGGTGTT CGAAGAAGAG ATCAATGGTG AAAAAATTGAC TGAATCATA      360
274 AATACTAGAC ATCAAAACGT GAAATACTTG CCTGGCATCA CTCTACCCGA CAATTTGGTT      420
276 GCTAATCCAG ACTTGATTGA TTCAGTCAAG GATGTCGACA TCATCGTTTT CAACATTCCA      480
278 CATCAATTTT TGCCCGTAT CTGTAGCCAA TTGAAAGGTC ATGTTGATTG ACACGTCAGA      540
280 GCTATCTCCT GTCTAAAGGG TTTTGAAGTT GGTGCTAAAG GTGTCCAATT GCTATCCTCT      600
282 TACATCACTG AGGAACTAGG TATTCATGT GGTGCTCTAT CTGGTGCTAA CATTGCCACC      660
284 GAAGTCGCTC AAGAACTAGT GTCTGAAACA ACAGTTGCTT ACCACATTCC AAAGGATTTC      720
286 AGAGGCGAGG GCAAGGACGT CGACCATAAG GTTCTAAAGG CCTTGTTCCA CAGACCTTAC      780
288 TTCCACGTTA GTGTCACTGA AGATGTTGCT GGTATCTCCA TCTGTGGTGC TTGAAGAAC      840
290 GTTGTTCGCT TAGGTTGTGG TTTGCTGCAA GGTCTAGGCT GGGGTAACAA CGCTTCTGCT      900
292 GCCATCCAAA GAGTCGGTTT GGGTGAGATC ATCAGATTCT GTCAAAATGTT TTTCCAGAAA      960
294 TCTAGAGAAG AAACATACTA CCAAGAGTCT GCTGGTGTG CTGATTGAT CACCACCTGC      1020
296 GCTGGTGGTA GAAACGTCAA GGTGCTAGG CTAATGGCTA CTTCTGGTAA GGACGCTGG      1080
298 GAATGTGAAA AGGAGTTGTT GAATGGCCAA TCCGCTCAAG GTTTAATTAC CTGCAAGAAA      1140
300 GTTACGAAAT GGTGGAAC ATGTGGCTCT GTCGAAGACT TCCCATTAT TGAAGCCGTA      1200
302 TACCAATCG TTTACAACAA CTACCCAATG AAGAACCTGC CGGACATGAT TGAAGAATTA      1260
304 GATCTACATG AAGATTAGAT TTATTGGAGA AAGATAACAT ATCATACTTC CCCCACTTT      1320
306 TTCGAGGCTC TTCTATATCA TATTCATAAA TTAGCATTAT GTCATTCTC ATAACACTTT      1380
309 (2) INFORMATION FOR SEQ ID NO: 6:
311      (i) SEQUENCE CHARACTERISTICS:
312      (A) LENGTH: 2946 base pairs
313      (B) TYPE: nucleic acid
314      (C) STRANDEDNESS: single
315      (D) TOPOLOGY: linear
317      (ii) MOLECULE TYPE: DNA (genomic)
319      (vi) ORIGINAL SOURCE:
320      (A) ORGANISM: GPD2
C--> 322      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
324 GAATTCGAGC CTGAAGTGCT GATTACCTTC AGGTAGACTT CATCTTGACC CATCAACCCC      60
326 AGCGTCAATC CTGCAAAATC ACCACCCAGC AGCACTAGGA TGATAGAGAT AATATAGTAC      120
328 GTGGTAACGC TTGCCTCATC ACCTACGCTA TGGCCGGAAT CGGCAACATC CCTAGAATTG      180
330 AGTACGTGTG ATCCGGATAA CAACGGCAGT GAATATATCT TCGGTATCGT AAAGATGTGA      240
332 TATAAGATGA TGTATACCCA ATGAGGAGCG CCTGATCGTG ACCTAGACCT TAGTGGCAAA      300
334 AACGACATAT CTATTATAGT GGGGAGAGTT TCGTGCAAAAT AACAGACGCA GCAGCAAGTA      360
336 ACTGTGACGA TATCAACTCT TTTTATTATTA TGTAATAAGC AAACAAGCAC GAATGGGGAA      420
338 AGCCTATGTG CAATCACCAA GGTGCTCCCT TTTTCCCAT TTGCTAATTT AGAATTTAAA      480
340 GAAACCAAAA GAATGAAGAA AGAAACAAA TACTAGCCCT AACCTGACT TCGTTTCTAT      540

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/308,207  
 DATE: 10/19/2000  
 TIME: 08:34:10

Input Set : A:\ES.txt  
 Output Set: N:\CRF3\10192000\I308207.raw

```

342 GATAATACCC TGCTTTAATG AACGGTATGC CCTAGGGTAT ATCTCACTCT GTACGTTACA 600
344 AACTCCGGTT ATTTTATCGG AACATCCGAG CACCCGCGCC TTCTCAACC CAGGCACCGC 660
346 CCCAGGTAAC CGTGCGGAT GAGCTAATCC TGAGCCATCA CCCACCCAC CCGTTGATGA 720
348 CAGCAATTCG GGAGGGCGAA AATAAACTG GAGCAAGGAA TTACCATCAC CGTCACCATC 780
350 ACCATCATAT CGCCTTAGCC TCTAGCCATA GCCATCATGC AAGCGTGTAT CTTCTAAGAT 840
352 TCAGTCATCA TCATTACCGA GTTTGTTTTT CTTCACATGA TGAAGAAGGT TTGAGTATGC 900
354 TCGAAACAAT AAGACGACGA TGGCTCTGCC ATTGGTTATA TTACGCTTTT GCGGCGAGGT 960
356 GCCGATGGGT TGCTGAGGGG AAGAGTGTTT AGCTTACGGA CCTATTGCCA TTGTTATTCC 1020
358 GATTAATCTA TTGTTACGCA GCTCTTCTCT ACCCTGTCAT TCTAGTATTT TTTTTTTTTT 1080
360 TTTTGGTTT TACTTTTTTT TCTTCTTGCC TTTTCTTCTT GTTACTTTTT TTCTAGTTTT 1140
362 TTTTCTTCC ACTAAGCTTT TTCTTGATT TATCCTTGGG TTCTTCTTTC TACTCCTTTA 1200
364 GATTTTTTTT TTATATATTA ATTTTAAAGT TTATGTATTT TGGTAGATTC AATTCTCTTT 1260
366 CCCTTTCCCT TTCTTCCGCT CCCCTTCCCT ATCAATGCTT GCTGTGAGAA GATTAACAAG 1320
368 ATACACATTC CTTAAGCGAA CGCATCCGGT GTTATATACT CGTCGTGCAT ATAAAAATTT 1380
370 GCCTTCAAGA TCTACTTTCC TAAGAAGATC ATTATTACAA ACACAACTGC ACTCAAAGAT 1440
372 GACTGCTCAT ACTAATATCA AACAGCACAA ACACGTGTCAT GAGGACCATC CTATCAGAA 1500
374 ATCGGACTCT GCCGTGTCAA TTGTACATTT GAAACGTGCG CCCTTCAAGG TTACAGTGAT 1560
376 TGGTTCTGGT AACTGGGGGA CCACCATCGC CAAAGTCATT GCGGAAAACA CAGAATTGCA 1620
378 TTCCCATATC TTCGAGCCAG AGGTGAGAAAT GTGGGTTTTT GATGAAAAGA TCGGCGACGA 1680
380 AAATCTGACG GATATCATAA ATACAAGACA CCAGAACGTT AAATATCTAC CCAATATTGA 1740
382 CTGCCCCCAT AATCTAGTGG CCGATCCTGA TCTTTTACAC TCCATCAAGG GTGCTGACAT 1800
384 CCTTGTTTTT AACATCCCTC ATCAATTTTT ACCAAACATA GTCAAACAAT TGCAAGGCCA 1860
386 CGTGGCCCCC CATGTAAGGG CCATCTCGTG TCTAAAAGGG TTGAGTTGG GCTCCAAGGG 1920
388 TGTGCAATTG CTATCCCTCT ATGTTACTGA TGAGTTAGGA ATCCAATGTG GCGCACTATC 1980
390 TGGTGCAAAC TTGGCACCAG AAGTGGCCAA GGAGCATTGG TCCGAAACCA CCGTGGCTTA 2040
392 CCAACTACCA AAGGATTATC AAGGTGATGG CAAGGATGTA GATCATAAGA TTTTGAAATT 2100
394 GCTGTTCCAC AGACCTTACT TCCACGTCAA TGTCATCGAT GATGTTGCTG GTATATCCAT 2160
396 TGCCGGTGCC TTGAAGAACG TCGTGGCACT TGCATGTGGT TTCGTAGAAG GTATGGGATG 2220
398 GGGTAACAAT GCCTCCCGAG CCATTCAAAG GCTGGGTTTA GGTGAAATTA TCAAGTTCCG 2280
400 TAGAATGTTT TTCCCAGAA CCAAAGTCGA GACCTACTAT CAAGAATCCG CTGGTGTGTC 2340
402 AGATCTGATC ACCACCTGCT CAGGCGGTAG AAACGTCAAG GTTGCCACAT ACATGGCCAA 2400
404 GACCGGTAAG TCAGCCTTGG AAGCAGAAAA GGAATTGCTT AACGGTCAAT CCGCCCCAAG 2460
406 GATAATCACA TGCAGAGAAG TTCACGAGTG GCTACAAACA TGTGAGTTGA CCCAAGAATT 2520
408 CCAATTATT CGAGGCAGTC TACCAGATAG TCTACAACAA CGTCCGCATG GAAGACCTAC 2580
410 CGGAGATGAT TGAAGAGCTA GACATCGATG ACGAATAGAC ACTCTCCCCC CCCCTCCCCC 2640
412 TCTGATCTTT CTGTTGCCT CTTTTTCCCC CAACCAATTT ATCATTATAC ACAAGTTCTA 2700
414 CAACTACTAC TAGTAACATT ACTACAGTTA TTATAATTTT CTATTCTCTT TTTCTTTAAG 2760
416 AATCTATCAT TAACGTTAAT TTCTATATAT ACATAACTAC CATTATACAC GCTATTATCG 2820
418 TTTACATATC ACATCACCGT TAATGAAAGA TACGACACCC TGTACACTAA CACAATTAAA 2880
420 TAATCGCCAT AACCTTTTCT GTTATCTATA GCCCTTAAAG CTGTTTCTTC GAGCTTTTCA 2940
422 CTGCAG 2946

```

```

424 (2) INFORMATION FOR SEQ ID NO: 7:
426 (i) SEQUENCE CHARACTERISTICS:
427 (A) LENGTH: 3178 base pairs
428 (B) TYPE: nucleic acid
429 (C) STRANDEDNESS: single
430 (D) TOPOLOGY: linear
432 (ii) MOLECULE TYPE: DNA (genomic)
434 (vi) ORIGINAL SOURCE:

```

## VERIFICATION SUMMARY

DATE: 10/19/2000

PATENT APPLICATION: US/09/308,207

TIME: 08:34:11

Input Set : A:\ES.txt

Output Set: N:\CRF3\10192000\I308207.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:35 M:220 C: Keyword misspelled or invalid format, (A) APPLICATION NUMBER:  
L:36 M:220 C: Keyword misspelled or invalid format, (B) FILING DATE:  
L:68 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:139 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:174 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:205 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:260 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:322 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:437 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:558 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:601 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:642 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:741 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:831 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:918 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1050 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1131 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1242 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1359 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1422 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1572 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:1990 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2006 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2020 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2034 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2048 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2062 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2076 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2090 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2104 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2118 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2132 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2146 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2160 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2177 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2243 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2363 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2417 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2459 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2546 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2560 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2574 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2588 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2602 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2616 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2630 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]

## VERIFICATION SUMMARY

DATE: 10/19/2000

PATENT APPLICATION: US/09/308,207

TIME: 08:34:11

Input Set : A:\ES.txt

Output Set: N:\CRF3\10192000\I308207.raw

L:2644 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2659 M:220 C: Keyword misspelled or invalid format, [(xi) SEQUENCE DESCRIPTION: SEQ ID NO:]  
L:2832 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=59, Value=[None]  
L:2958 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=60, Value=[None]  
L:2997 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=61, Value=[None]  
L:3036 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=62, Value=[None]  
L:3072 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=63, Value=[None]  
L:3108 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=64, Value=[None]  
L:3153 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=65, Value=[None]  
L:3272 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=67, Value=[None]  
L:3467 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:68